



# Using Regions of Interest for Adaptive Image Retrieval

Michael Springmann, Heiko Schuldt

*University of Basel, Database and Information Systems Group  
Bernoullistr 16, CH-4056, Basel, Switzerland  
{michael.springmann, heiko.schuldt}@unibas.ch*



The work has been supported by the Swiss National Science Foundation (SNF) in the context of the project Query by Sketching under contract No. 200021-117800 / 1

# Content-Based Image Retrieval

Query by Example



Query Image

Feature Vector

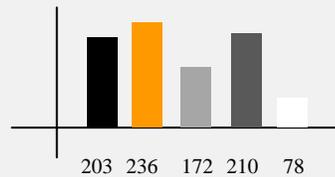
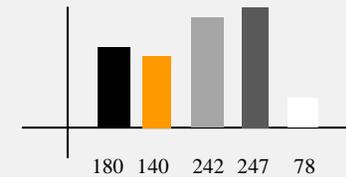


Image Collection



Feature Vector



$$\sum_{i=0}^d (a_i - b_i)^2$$



# Simplified Comparison with Text IR

	CBIR	Text IR
What is the query?	Feature(s) of Image(s)	Set of keywords
What is desired result?	Ranked list of most-similar Images	(Ranked) list of Documents containing keywords
What contributes to distance/similarity?	Every difference between feature vectors (query + compared image)	Only keywords found in documents
How is query refined (e.g. using relevance feedback)	Adjust weights of features / images, add new features/images	Add new keywords, adjust weights

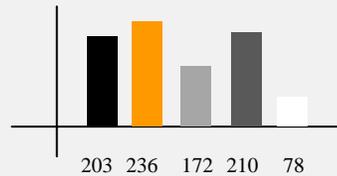
# Adaptability in Query by Example

What contributes to distance?

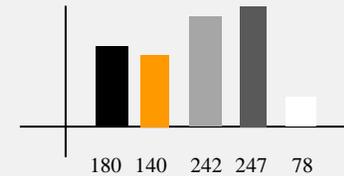


Query Image

Feature Vector



Feature Vector



$$\sum_{i=0}^d (a_i - b_i)^2$$

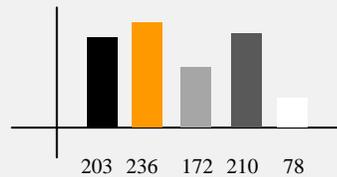
# Adaptability in Query by Example

What contributes to distance?

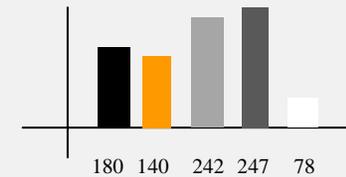


Query Image

Feature Vector



Feature Vector



$$\sum_{i=0}^d (a_i - b_i)^2$$



# Overview of the Approach

Selecting Regions of Interest (ROI)

Matching Regions in Images

Similarity Between Corresponding  
Regions

Conclusion & Outlook

Relevance Feedback



# Selecting Regions of Interest (ROI)

- Traditional Approaches: Rectangular Bounding Box



Use just the bbox as new query image

... but this matches entire images,  
not just interesting regions.

# Selecting Regions of Interest (ROI)

- Traditional Approaches: Automatic Segmentation

Blobworld:




blob and feature importance:					
	blob (overall)	color	texture	location	shape
blob 2	very	not	not	not	very

Querying from 35000 images (2000 returned by the filter).




1: 77045 (score = 0.97)
[New query](#)




2: 107090 (score = 0.96)
[New query](#)



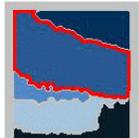

3: 374088 (score = 0.96)
[New query](#)



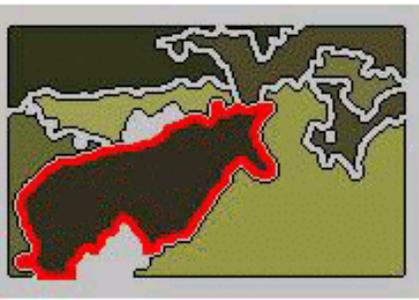

4: 228093 (score = 0.96)
[New query](#)




5: 205042 (score = 0.96)
[New query](#)

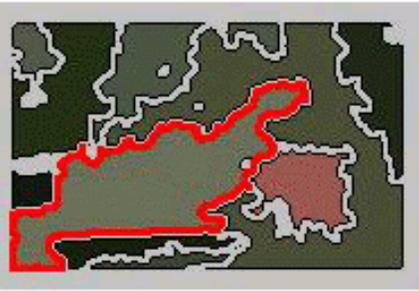



6: 26052 (score = 0.96)
[New query](#)







5: 205042 (score = 0.96)
[New query](#)



# Selecting Regions of Interest (ROI)

- Intrinsic Bounds to Automatic Image Segmentation





# Selecting Regions of Interest (ROI)

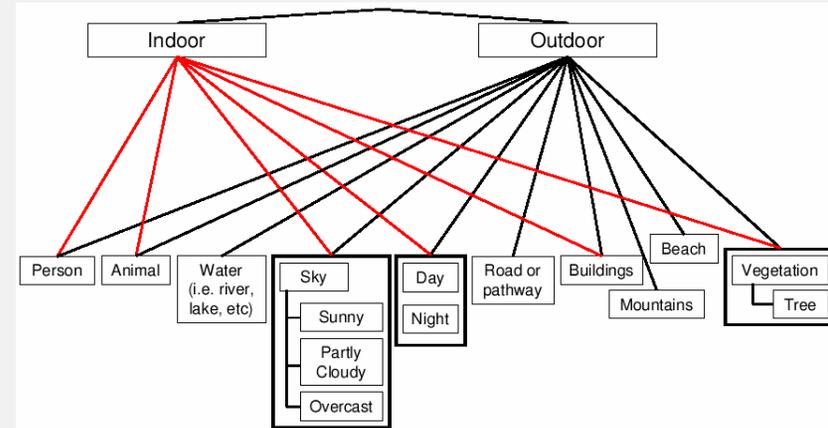
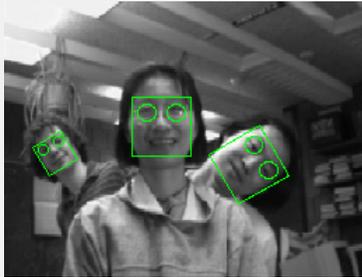
## Using Interactive Paper or Tablet PC



Used for query definition *and* relevance feedback

# Identify High-Level Concepts

- Face Detection
- ImageCLEF
  - Visual Concept Detection Task
  - Automatic Medical Annotation



- ALIPR

ALIPR

Keyword(s):

**Top 15 Computer-Predicted Tags**  
ALIPR is like a child trying to learn about the world. Please help us to teach ALIPR. Check those correctly annotated words.

building  historical  people  landscape  grass  
 animal  face  food  wild\_life  rock  
 snow  fox  water  beach  primate

Thought of other terms missed by ALIPR? Please add here, separated by commas !!  
  and make the picture searchable

Optional information:  
Picture title:   
URL to see related pictures:   
Copyright (hypertext ok):

© alipr.com 2005-2007 Patent Pending. All rights reserved. Do NOT upload inappropriate images. Pictures may be subject to copyright!



# Selecting Regions of Interest - Summary

- User selects Region of Interest (ROI)
- Input devices more appropriate than mouse & keyboard
- System tries to find high-level concepts within ROI

But how to match selected region?



# Overview of the Approach

Selecting Regions of Interest (ROI)

Matching Regions in Images

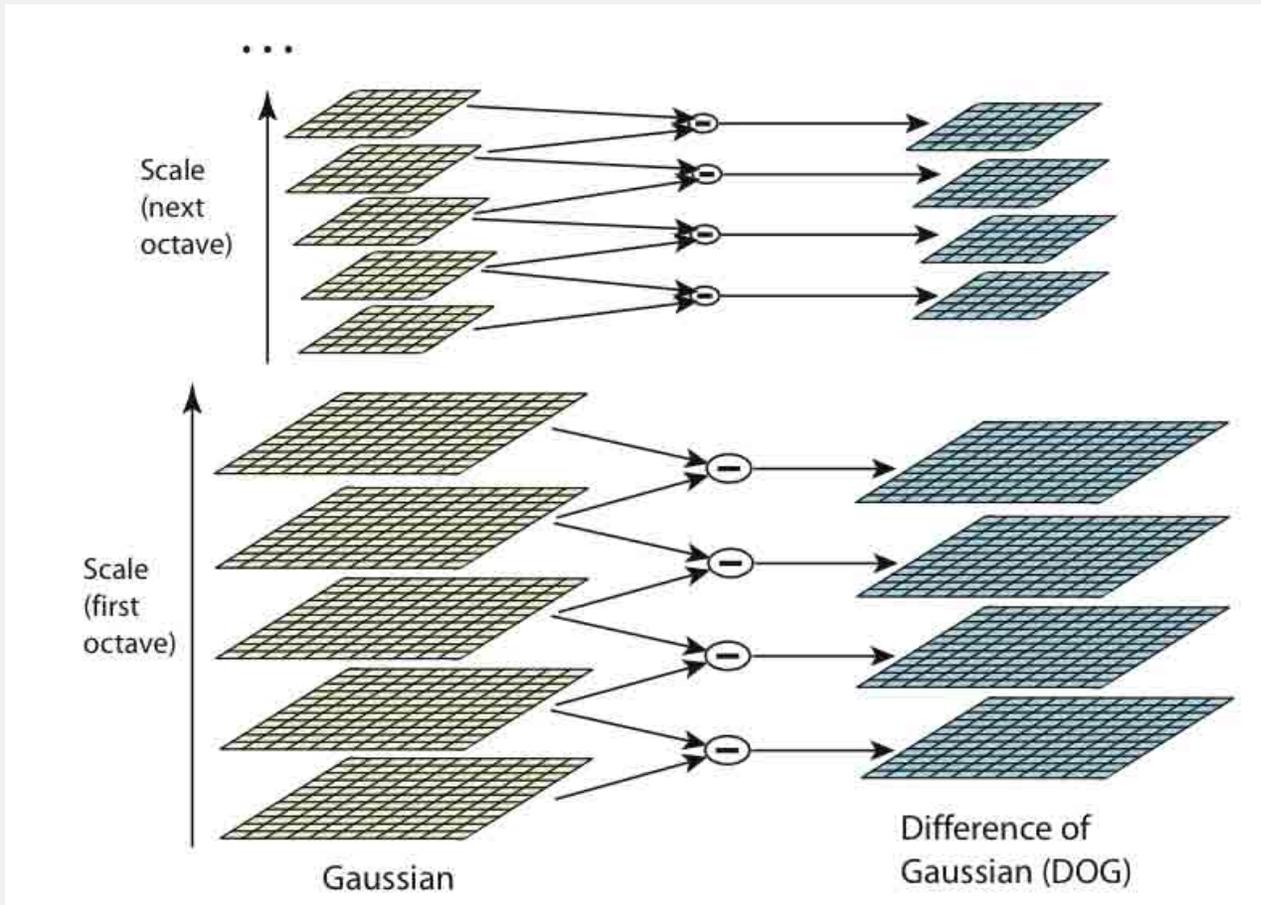
Similarity Between Corresponding  
Regions

Relevance Feedback

Conclusion & Outlook

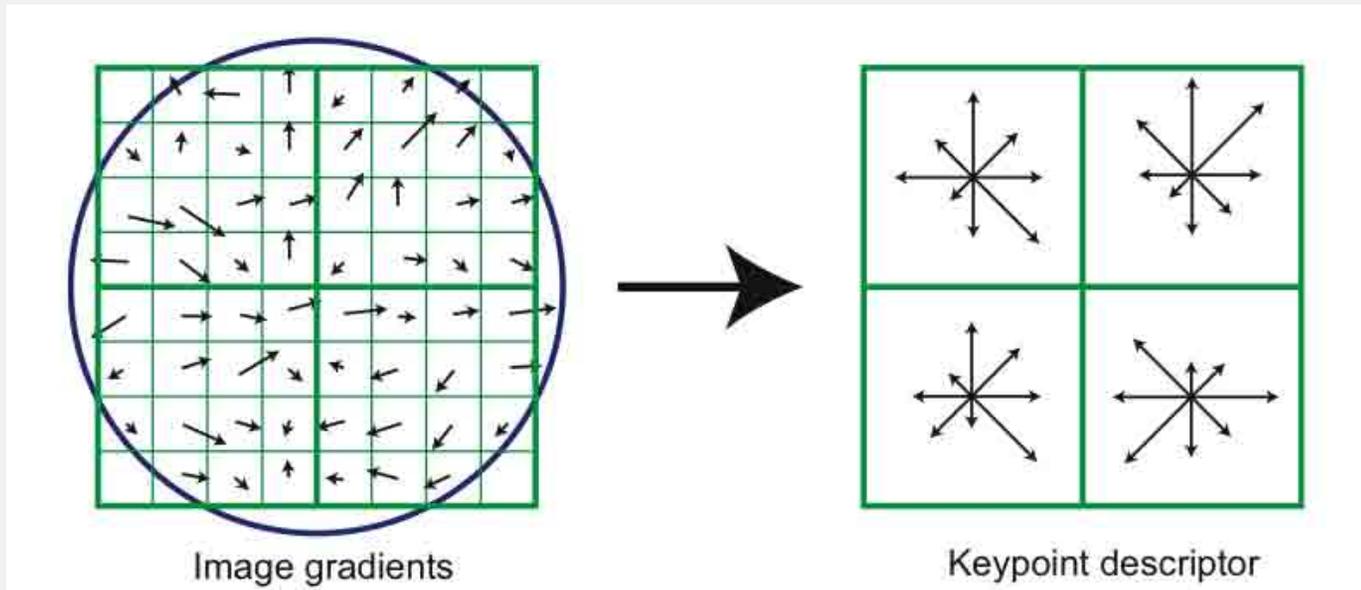
# Matching Regions in Images: Keypoints

Generic Solution: SIFT (Scale-Invariant Feature Transform)



# Matching Regions in Images: Keypoints

Generic Solution: SIFT (Scale-Invariant Feature Transform)





# Matching Regions in Images: Keypoints

Generic Solution: SIFT (Scale-Invariant Feature Transform)





# Keypoint Matching



1481 Keypoints

2536 Keypoints

For each keypoint in query

Many!

find the two most similar keypoint in reference

128 dim dist.  
- Slow!

determine ratio:  $\text{best\_distance} / \text{second\_distance}$

if better than threshold, keep as match

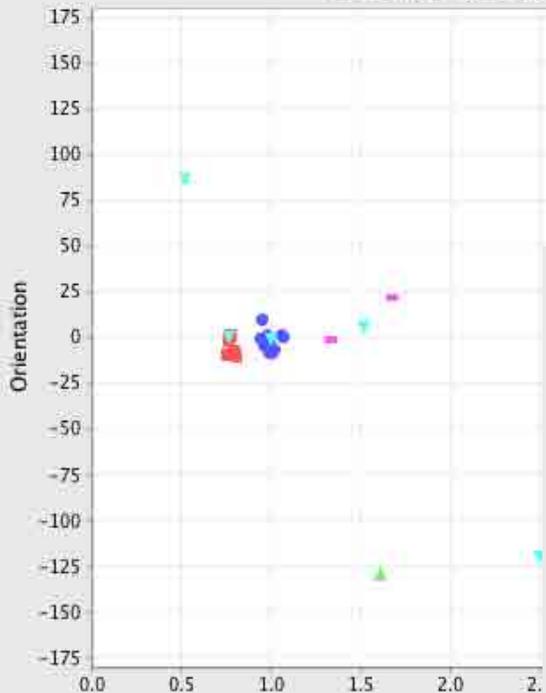
End

Hard to define!



# Clustering and Filtering

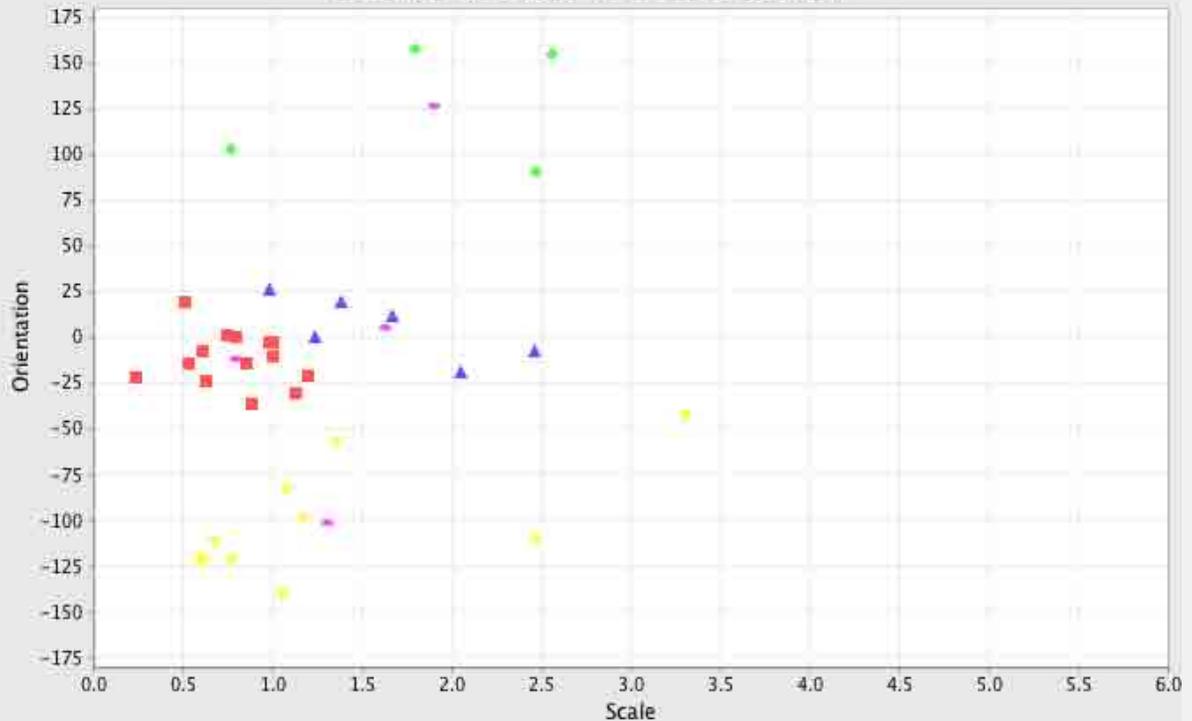
Estimated Scale and Orientation



Input:  
List of 50 best SIFT  
matches (no threshold)

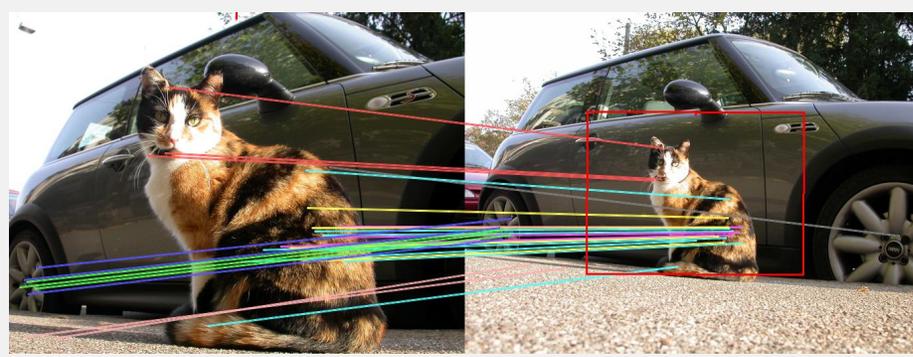
Filter Outliers with  
RANSAC

Estimated Scale and Orientation





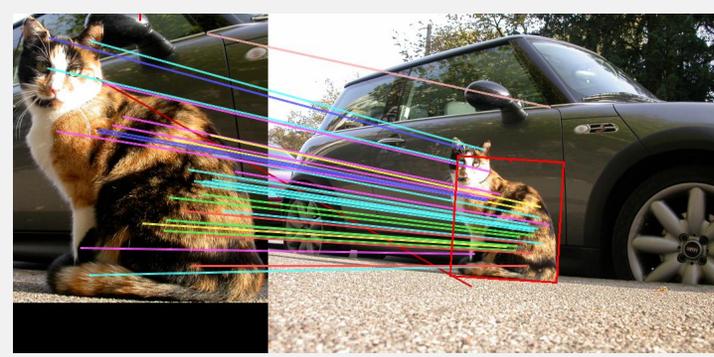
# Keypoint Matching



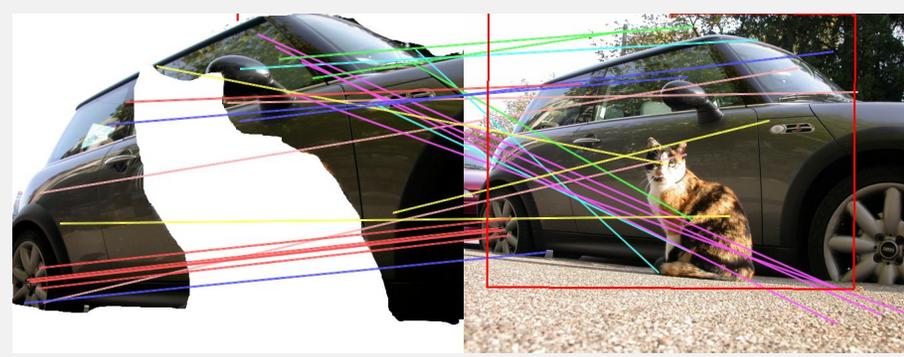
1481 Query keypoint: 924 ms  
Early Termination: 629 ms



684 Query keypoint: 438 ms  
Early Termination: 311 ms



900 Query keypoint: 576 ms  
Early Termination: 401 ms



536 Query keypoint: 350 ms  
Early Termination: 236 ms



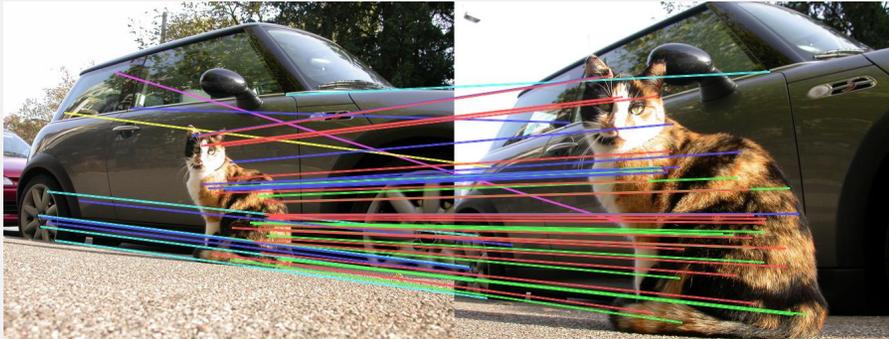
# Further Optimizations possible

- Did not use any index so far
- Can be performed in parallel using multiple threads / computing cluster / Grid
- Add constraints:
  - Allow only certain rotations and scaling
  - Can be evaluated based on keypoint metadata
  - Example: Allow only  $10^\circ$  of rotation  $\rightarrow$  152 ms for cat region
- Remember:
  - Number of images to perform matching on can be filtered based on high-level metadata (face, visual concept, tags,...)



# Feedback

- Can be traced back to originating keypoint



- Can be used to refine transformation constraints
- Can be used to refine high-level concepts



# Matching Regions - Summary

- Use invariant keypoint descriptors like SIFT originating from object recognition domain
- Allow more matches, but add clustering & filtering
- Use Early Termination and constraints to reduce time

But how to determine similarity?



# Overview of the Approach

Selecting Regions of Interest (ROI)

Matching Regions in Images

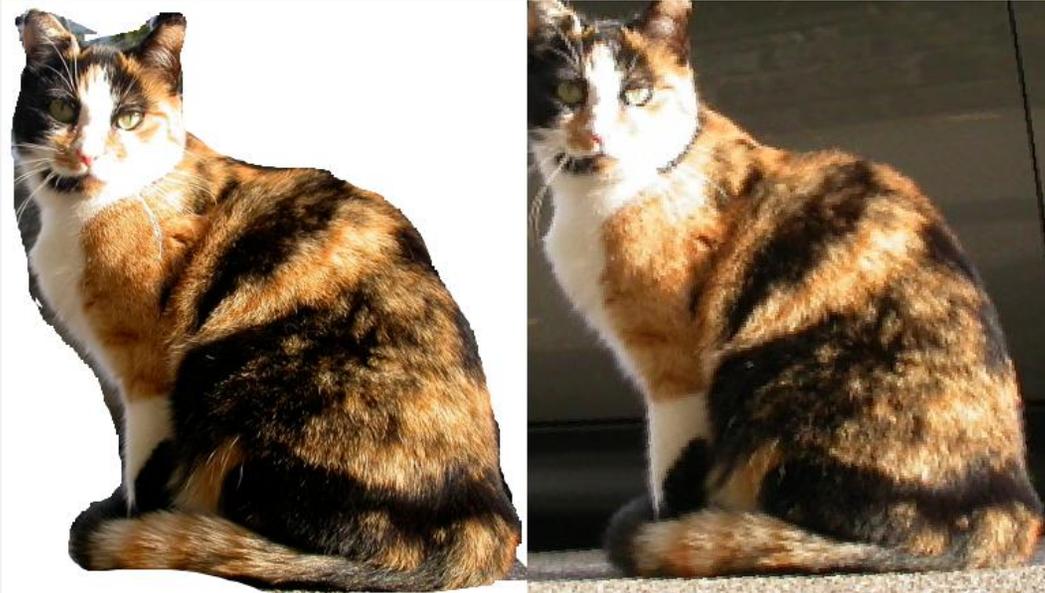
Similarity Between Corresponding  
Regions

Conclusion & Outlook

Relevance Feedback



# Determine and Apply Affine Transformation





# Similarity Between Corresponding Regions

## Image Distortion Model (IDM)

Ignore Pixels outside Query Region

Local Context

Warp range

$$d_{IDM}(A, B) = \sqrt{\sum_{x=1}^{width} \sum_{y=1}^{height} \min_{i=x-w}^{x+w} \min_{j=y-w}^{y+w} (A_{i,j} - B_{i,j})^2}$$



# Conclusion

## Selecting Regions of Interest (ROI)

using Interactive Paper or Tablet PCs to

- Give maximum flexibility to the user
- Better understand what is the need (concepts)

## Matching Regions in Images

using keypoints within query regions

- use clustering and filtering
- determine transformation

## Similarity Between Corresponding Regions

Based on Image Distortion Model

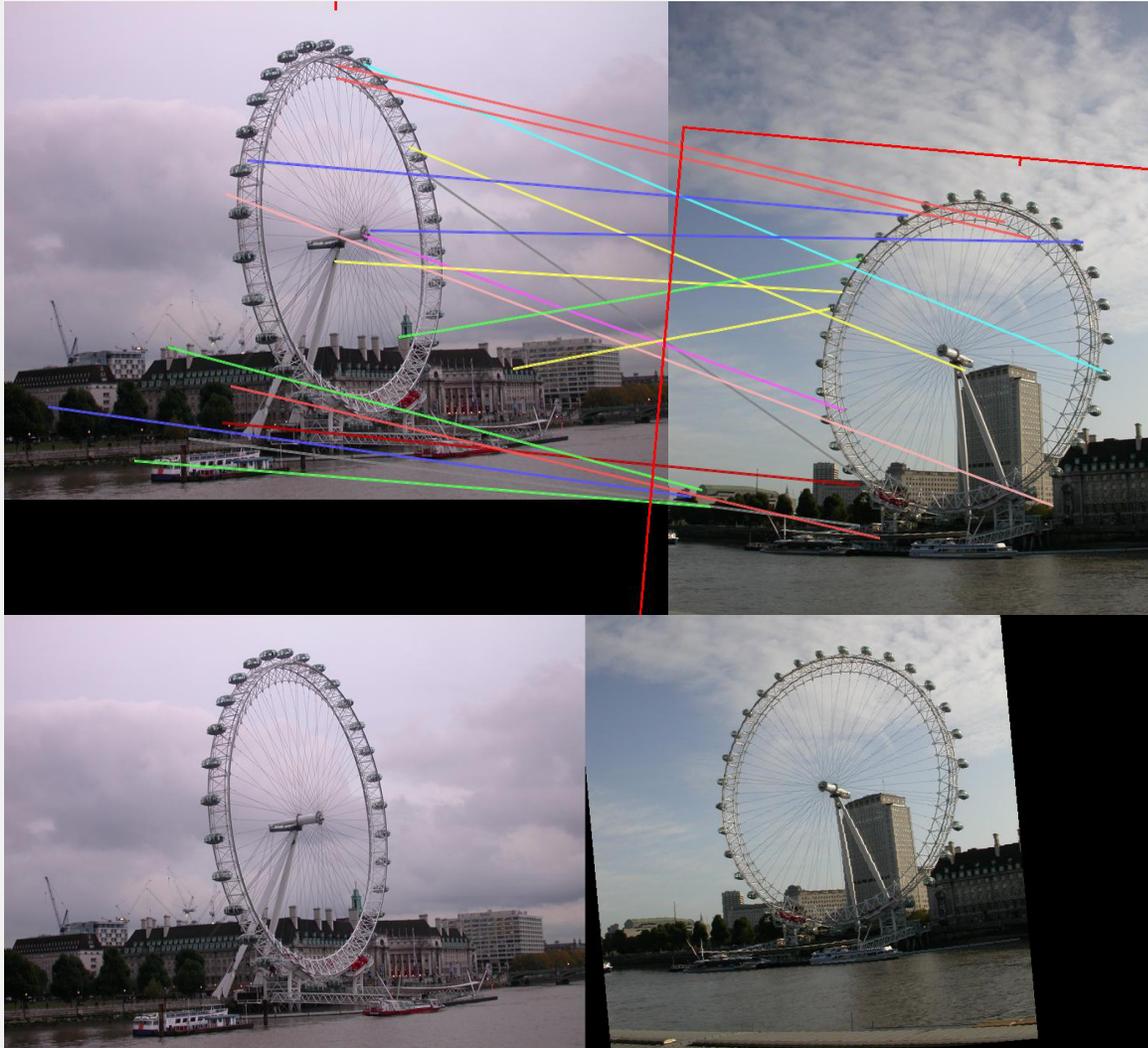
- Ignore pixels outside query regions

Relevance Feedback

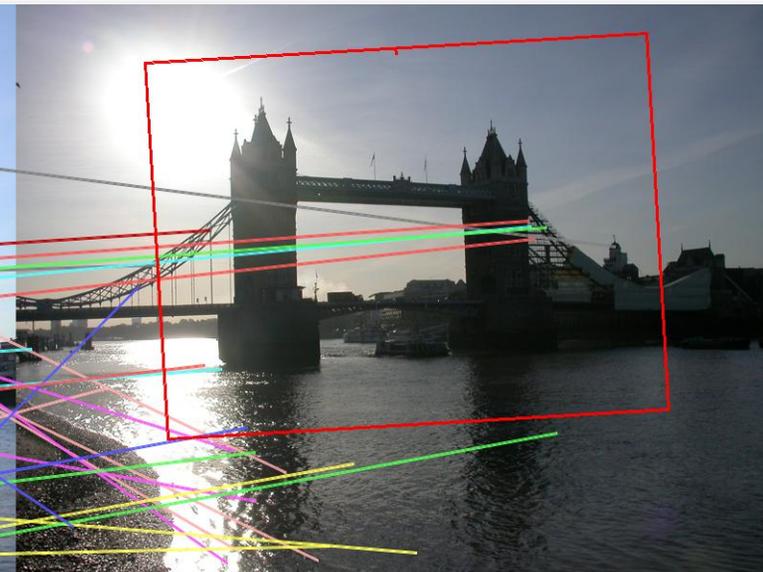
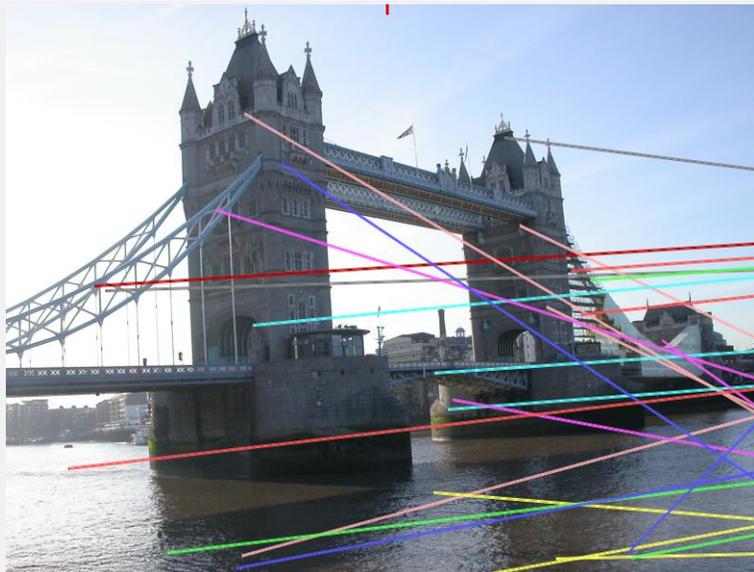
Whenever useful to adapt to user



# Examples – London Eye

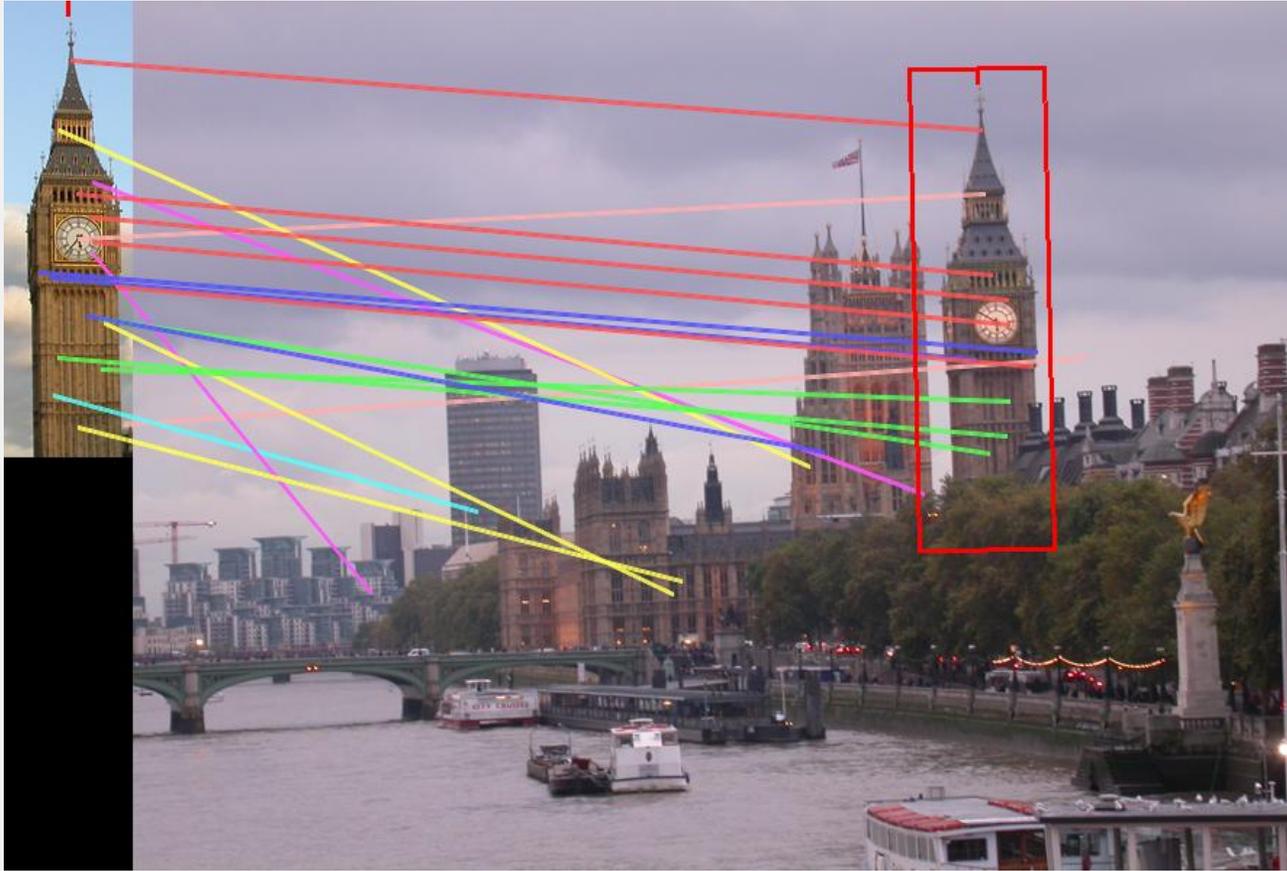


# Examples – Tower Bridge





# Examples – Big Ben



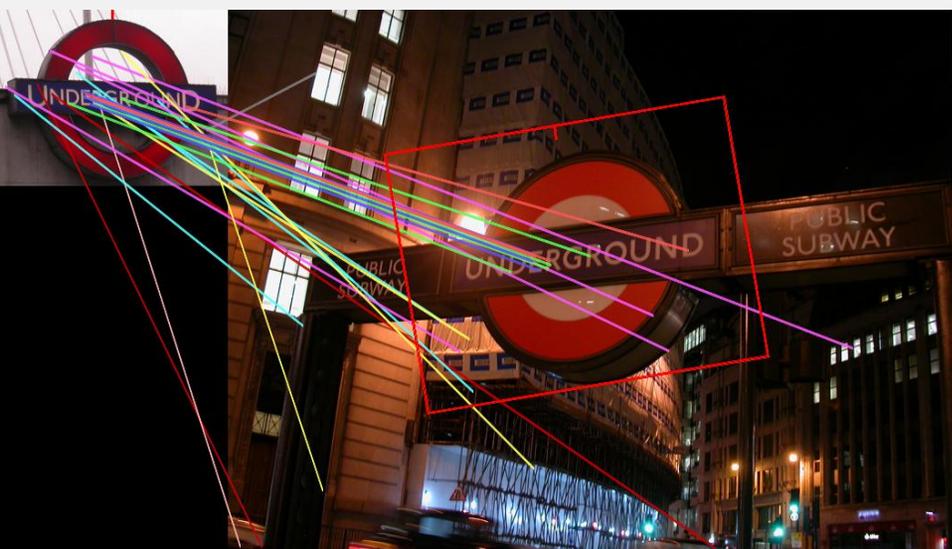


# Examples – Phone Booth

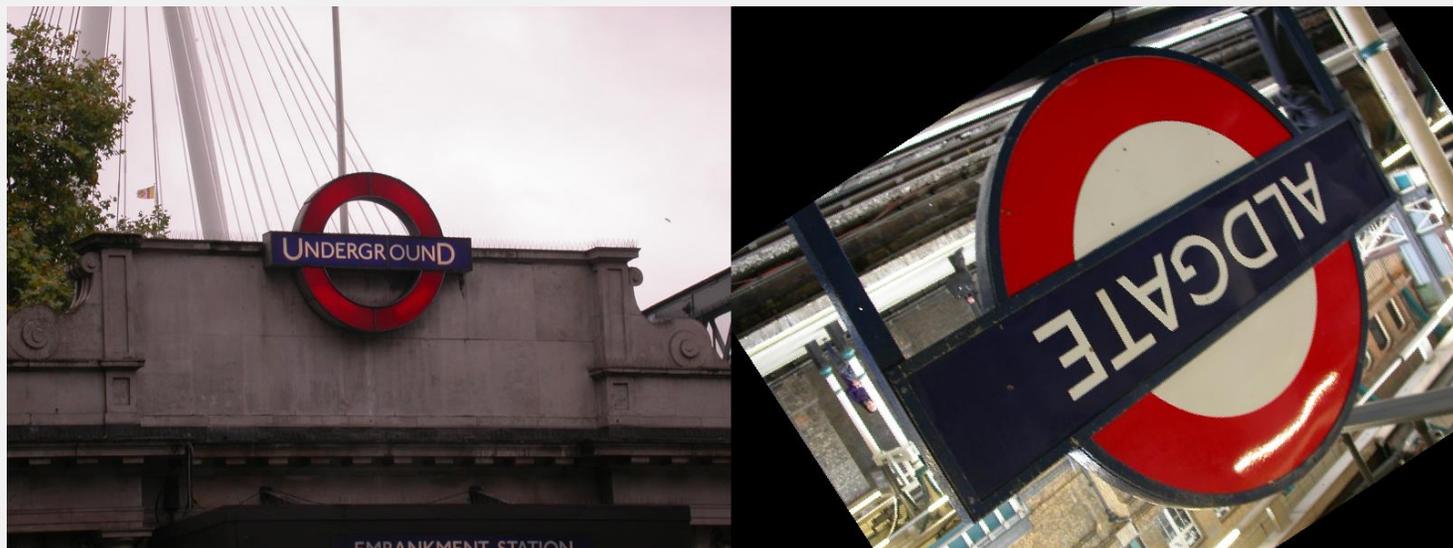




# Examples – Underground



# Examples – Underground (2)





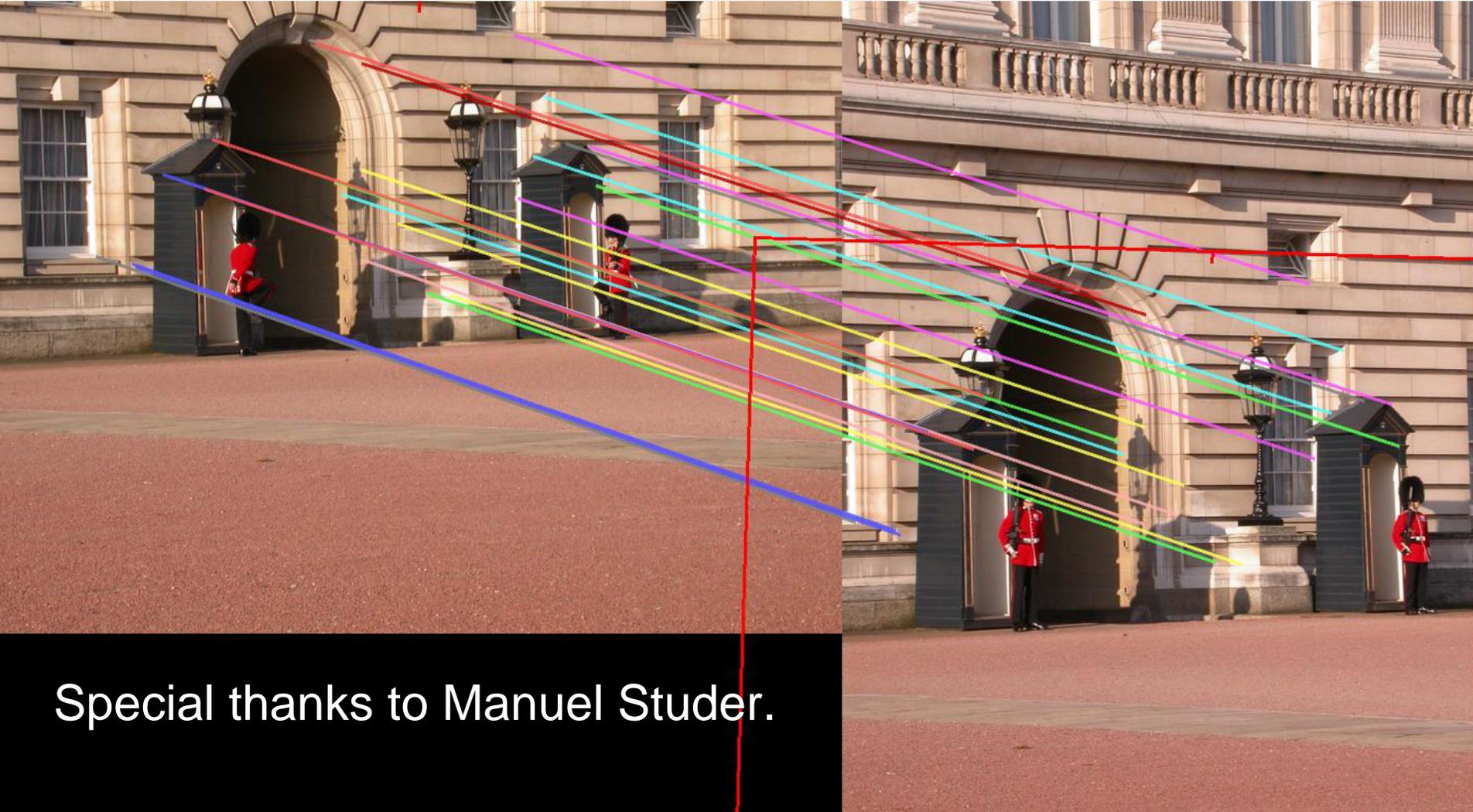
# Outlook

- Work on Keypoints
  - SIFT is ideal for all images
  - Use additional color information in color images
  - Edge-based keypoint descriptors required for sketches
- Combinations of regions
  - Semantics and implementation of AND, OR, AVERAGE
- Evaluate approach
- Distributed computation to reduce execution time



# Thanks!

Thank you for your attention.



Special thanks to Manuel Studer.